DOCUMENT RESUME

ED 392 843 TM 024 686

AUTHOR	Meshbane, Alice; Morris, John D.
TITLE	Assuming Equal vs. Unequal Prior Probabilities of Group Membership in Discriminant Analysis: Effect on Predictive Accuracy.
PUB DATE	Apr 95
NOTE	20p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Classification; Comparative Analysis; *Discriminant Analysis; *Group Membership; *Prediction; *Probability; *Sample Size
IDENTIFIERS	Accuracy; Bootstrap Methods

ABSTRACT

Cross-validated classification accuracies were compared under assumptions of equal and varying degrees of unequal prior probabilities of group membership for 24 bootstrap and 48 simulated data sets. The data sets varied in sample size, number of predictors, relative group size, and degree of group separation. Total-group hit rates were used to compare the relative accuracies across six assumptions about prior probabilities. Contrary to expectations, use of population priors did not always yield the highest hit rate. When group sizes were similar, equal priors yielded greater classification accuracy than sample estimated priors. Results suggest that, when group sizes are similar, use of unequal priors may lead to a decrement in classification accuracy, even with knowledge of population priors. (Contains 5 tables and 13 references.) (Author)

Assuming Equal vs. Unequal Prior Probabilities of Group Membership in Discriminant Analysis: Effect on Predictive Accuracy

Alice Meshbane and John D. Morris
Florida Atlantic University

Paper presented at the annual meeting of the American Educational Research Association, April 1995, San Francisco, CA

Assuming Equal vs. Unequal Prior Probabilities of Group Membership in Discriminant Analysis: Effect on Predictive Accuracy

ABSTRACT. Cross-validated classification accuracies were compared under assumptions of equal and varying degrees of unequal prior probabilities of group membership for 24 bootstrap and 48 simulated data sets. The data sets varied in sample size, number of predictors, relative group size, and degree of group separation. Total-group hit rates were used to compare the relative accuracies across six assumptions about prior probabilities. Contrary to expectations, use of population priors did not always yield the highest hit rate. When group sizes were similar, equal priors yielded greater classification accuracy than sample estimated priors. Results suggest that, when group sizes are similar, use of unequal priors may lead to a decrement in classification accuracy, even with knowledge of population priors.

Theoretically, the assumption of unequal prior probabilities should lead to higher cross-validated classification accuracy as the difference between population group sizes increases. "Where we might tend to oversupply small groups and undersupply large ones by using resemblance as the sole basis for classification we introduce a corrective effect by taking prior probabilities of group membership into account" (Tatsuoka, 1988, p. 360). Consistent with this expectation, Rudolph and Karson (1988) found that estimated error rates using population priors were consistently lower than estimated error rates using equal priors.

Although classification accuracy should increase with knowledge and use of population group sizes, these values are rarely known. Consequently, sample estimated values are generally used. However, the use of sample estimated values may be unwise. Huberty (1994, p. 65) argues that "priors should *not* correspond to the relative sample sizes unless ... a proportional sampling plan was utilized." Of course, proportional sampling presumes knowledge of population priors. Lindeman, Merenda, and Gold (1980, p. 211) point out that, "in most practical applications, the values of the prior probabilities are not known with sufficient accuracy to justify their use." Hence, these researchers have urged caution in using anything but equal prior probabilities of group membership for classification.

The purpose of this study was to compare assumptions of equal versus varying degrees of unequal prior probabilities of group membership on cross-validated classification accuracy. The goal was to get some idea of the degree of difference in accuracy we might expect on application of these assumptions in practical classification problems. Implicit in this goal is the question of whether the increment

that may be afforded by assuming unequal priors is worth the risk when population priors are unknown.

Method

Cross-validated classification accuracies were compared under a variety of bootstrap and simulated data conditions (different sample sizes, predictor counts, relative group sizes, and prior probability assumptions) for the two-group classification problem. A total of 24 bootstrap and 48 simulated data sets were considered for each of six assumptions about prior probabilities of group membership:

(1) sample n / sample N (Sample condition);
(2) 1 / number of groups (Equal condition);
(3) population n / population N (Pop+0 condition);
(4) group size for smaller group is 15% larger (Pop+.15 condition);
(5) group size for smaller group is 30% larger (Pop+.30 condition); and
(6) group size for smaller group is 45% larger (Pop+.45 condition).

The bootstrap data sets were obtained from 24 real data sets used in a prior classification methodology study (Morris & Huberty, 1987). No pathological distributional problems are known in any of the data sets; it is expected that they are much as one would find in typical classification studies.

The 48 simulated populations were constructed according to multivariate normal models, with N ranging from 1270 to 2000. The group means and covariance matrices needed for input to the population creation program were obtained from the 24 real data sets mentioned in the previous paragraph. For 24 populations, group sizes were set to 1000. The remaining 24 populations were identical to these, except that group sizes were proportional to the sample sizes found in the real data sets.

The process for creating a population manifesting a specified covariance matrix is described in Morris (1975). The random normal deviates required by this method were created using the "Rectangle-Wedge-Tail" method (Marsaglia, MacLaren, & Bray, 1964), with the required uniform random numbers generated by Park & Miller's (1988) "minimal standard" algorithm. A FORTRAN computer program (modified for 64-bit word MS FORTRAN 5.0) provided by Dolker and Halperin was used for the variable generation.

Classification rules for a randomly selected (with replacement) sample of the desired size were built with adjustments made for each of the six assumptions about prior probabilities. The adjusted classification rules were used to classify the entire

population according to Tatsuoka's (1988) minimum chi square rule. This procedure was repeated 1000 times for the 24 bootstrap data sets and 250 times for each of the 48 simulated populations, and the mean number of total-group correct classifications was used to compare the relative accuracies of the six assumptions.

In order to be more confident in the results of this simulation, and in accord with Knuth's (1969, p. 156) recommendation that "the most prudent policy for a person to follow is to run each Monte Carlo program at least twice using quite different sources of random numbers, before taking the answers of the program seriously," the entire simulation results were replicated. In the replication, Wichman and Hill's (1982) uniform random number generator was used. This algorithm generates uniform random numbers by a triple modulo method. As described in Wilkinson (1987, p. 34 of the DATA module), "each uniform is constructed from three multiplicative congruential generators with prime modulus," using 13579, 12345, and 313 as initial seeds. While there were some small differences in the results of the replication, none were systematic, none were judged of importance, and the implications were the same. These replication results are available on request; the results presented in this paper are from the first random number generation method mentioned.

Results

For each of the data sets, Tables 1, 2, and 3 give a short description, an index of group separation ($\underline{D}$), the number of cases in each group ($\underline{n}_1$ and $\underline{n}_2$), the number of predictor variables ($\underline{p}$), and a comparison of the cross validated classification performance for each assumption about prior probabilities. Tables 2 and 3 also include an index of disproportionality ($\underline{I}$), calculated as (($\underline{n}_{larger}$ / $\underline{n}_{smaller}$) * 100). The best performing assumptions are underlined. The difference in performance between underlined and nonunderlined assumptions was considered statistically and practically significant based on subjectively established criteria (α = .00001 plus a mean difference in hit rates of .002, which represents 4 hits for data sets with 2000 cases). The risk of a Type I error was actually much higher than .00001 due to the large number of significance tests conducted. Although statistical significance was considered less important than practical significance, an overall Hotelling T^2 test, and then pairwise post hoc comparisons (multivariate analog of the Scheffé post hoc test; see Morrison, 1976, p. 147-148 for a description) were used to contrast the classification hit rates for the six assumptions.

Results of Simulation for Data Sets with Equal Group Sizes (#1-24)

The Equal and Pop+0 assumptions, which yield identical results with equal group sizes, were expected to outperform the other four assumptions in all 24 data sets. As indicated in Table 1, the Equal and Pop+0 conditions were top contenders in all but one data set (#15), and yielded the highest (though not always significantly higher) hit rates in 18 data sets (#5 - 8, 10 - 14, 16 - 24). Thus, these assumptions were the best performers most of the time rather than all of the time, which was somewhat contrary to expectations.

Insert Table 1 About Here

The Sample assumption was expected to perform less well than the Equal and Pop+0 assumptions due to sampling error inherent in the random sampling process, but was still expected to outperform the three erroneous assumptions (Pop+.15, Pop+.30, Pop+.45). Results were consistent with this expectation. The Sample assumption was a top contender in the same 23 data sets as the Equal and Pop+0 assumptions. Nevertheless, compared to the Equal and Pop+0 assumptions, the Sample assumption yielded lower hit rates (though not significantly, based on $\alpha = .00001$) in 21 of the data sets (4-24).

The rank order of the erroneous assumptions was expected to be Pop+.15, Pop+.30, and Pop+.45 (i.e.,from least to most discrepant with actual group sizes). Results were consistent with this expectation. The Pop+.15 condition performed better than the other two erroneous assumptions and worse than the Equal, Pop+0, and Sample assumptions. The Pop+.15 assumption was a top contender in 12 of the 24 data sets, and was the best performer in two data sets (#9, 15). The Pop+.30 assumption significantly outperformed the Pop+.45 assumption in 19 of the 24 data sets (#6 - 24).

Results of Simulation for Data Sets with Group Sizes Proportional to Real Data Set Sizes (#25-48)

The Pop+0 assumption was expected to outperform the other five assumptions in all 24 data sets. As shown in Table 2, the Pop+0 condition was a top contender in all but four data sets (#39 - 42), and yielded the highest (though not always significantly higher) hit rates in 11 data sets (#30 - 34, 38, and 43 - 47). No other

assumption performed as well. Thus, although the Pop+0 was the best performer overall, the results were somewhat contrary to expectations because this assumption did not yield the highest hit rate with every data set.

Insert Table 2 About Here

The Sample assumption was expected to perform less well than the Pop+0 assumption, again due to sampling error, but to outperform the four other assumptions. Results were consistent with this expectation for the Pop+0, Pop+.15, Pop+.30, and Pop+.45 assumptions. Compared to the Pop+0 assumption, the Sample assumption yielded lower hit rates in 20 of the data sets (#27, 28, 30 - 47), though this difference was statistically significant only for data set #45. The Sample assumption was a top contender in 18 of the 19 data sets for which the Pop+0 assumption was also a top contender (#25 - 37, 39, 43, 44, and 46-48), and had the highest hit rates (though not significantly higher) in 2 data sets (#26 and 48). None of the erroneous assumptions matched this performance. Compared to the Pop+.15 condition, which was the best performing erroneous assumption, the Sample assumption yielded higher hit rates (though not always significantly higher) in 15 data sets (#26, 30 - 34, 36 - 38, 43 - 48).

In the 20 data sets with unequal group sizes, the Equal assumption worked better than the Sample assumption only in data sets with small differences between group sizes (#27, 29, 35 - 37, 40 - 42, 44, 45). In the seven data sets with an index of disproportionality greater than 129 (#30 - 32, 34, 38, 43, 47), the Sample assumption outperformed the Equal assumption. The Sample assumption also outperformed the Equal assumption in three data sets with smaller differences between group sizes (#33, 46, 48).

As with equal group sizes, the rank order of the erroneous assumptions with unequal group sizes was expected to be Pop+.15, Pop+.30, and Pop+.45 (i.e., from least to most discrepant with actual group sizes). Results were consistent with this expectation, parallel to the findings for equal group sizes. The Pop+.15 condition performed better than the Pop+.30 and Pop+.45 assumptions and worse than the Pop+0 and Sample assumptions. The Pop+.15 assumption was a top contender in 14 of the 24 data sets, was the best performer in five data sets (# 29, 39 - 42), and outperformed (though not always significantly) the Pop+.30 assumption in 20 data

sets (#29 - 48). The Pop+.30 assumption outperformed the Pop+.45 assumption, though not always significantly, in 21 of the data sets (#26 and 29 - 48).

Results for Bootstrap Data Sets (Data Sets 49-72)

Results for the bootstrap data sets were quite similar to results for the simulated data sets. As shown in Table 3, tne Pop+0 condition was a top contender in all data sets, and had the highest hit rate in 12 data sets (#54 - 56, 58, 61 - 62, and 67 - 72). Nevertheless, other assumptions yielded higher hit rates (though not always significantly higher) in eight data sets (#51, 53, 57, 59, 60, 64 - 66). Thus, as with the simulated data sets, these results were somewhat contrary to expectations because the Pop+0 assumption did not yield the highest hit rate with every data set.

Insert Table 3 About Here

The Sample assumption was a top contender in all but one data set (#70), and had the second highest hit rate (behind Pop+0) in nine data sets (#54 - 56, 58, 61, 62, 67, 71, 72). Compared to the Pop+0 assumption, the Sample assumption yielded lower hit rates in 22 of the data sets (#51 - 72), though this difference was statistically significant only for data set #70. Compared to the Pop+.15 condition, the Sample assumption yielded higher hit rates (though not always significantly higher) in 14 data sets (#52, 54, 55, 56, 58, 61 - 63, 67 - 72). Thus, the performance of the Sample assumption relative to the erroneous assumptions matches what was found in the simulated data sets.

The Equal assumption worked better than the Sample assumption only in data sets with similar group sizes (#51 - 53, 57, 59, 60, 63 - 66, 68 - 70). The Sample assumption outperformed the Equal assumption in the seven data sets with indices of disproportionality greater than 129 (#54 - 56, 58, 62, 67, 71), as well as in two data sets with more similar group sizes (#61, 72). The Pop+.15 condition performed better than the Pop+.30 and Pop+.45 assumptions and worse than the Pop+0 and Sample assumptions. The Pop+.15 assumption was a top contender in all but three data sets (#61, 63, 70), was the best performer in two data sets (57, 65), and outperformed (though not always significantly) the Pop+.30 assumption in 20 data sets (#52, 54 - 72). The Pop+.30 assumption outperformed the Pop+.45 assumption, though not always significantly, in 21 of the data sets (#52 - 67). Again, these bootstrap results were similar to what was found in the simulated data sets.

Discussion

Pop+0 vs. Other Assumptions

Although the Pop+0 assumption was the best performer in an absolute sense, its performance relative to the other five assumptions was not as good as predicted. The erroneous assumptions occasionally performed much better than would be expected based on their discrepancy with population sizes. For example, in some data sets with unequal population sizes, the erroneous assumption of equal priors yielded a higher hit rate than the correct assumption about population priors.

At first glance, these results appear to be inconsistent with Rudolph and Karson's (1988) finding of consistently lower error rate estimates using population priors rather than equal priors. This apparent inconsistency may be due to differences in relative population sizes between the two studies. In the Rudolph and Karson study, the population priors were .9 and .1, reflecting a large discrepancy in population sizes. In the present study, equal priors yielded a higher hit rate than population priors only in data sets with similar group sizes. In all data sets with non-trivial differences in group sizes (I greater than 129), use of population priors increased the hit rate over equal priors.

Further support for this explanation of the apparent inconsistency between the two studies comes from a partial replication of the simulation. For data sets with similar group sizes (I less than or equal to 129) in which an erroneous assumption outperformed the Pop+0 assumption, new simulated data sets were created, each with 900 1's and 100 2's. As in the Rudolph and Karson study, the Pop+0 assumption outperformed the erroneous assumptions for every data set. These results are displayed in Table 4. Thus, our findings were consistent with Rudolph and Karson for data sets with dissimilar group sizes.

Insert Table 4 About Here

Still, it may seem counterintuitive that in any data set, erroneous assumptions about priors could yield higher hit rates than the correct assumption. An explanation for this is related to the differential effectiveness of the two classification rules when different priors are used. Suppose the two classification rules are equally effective in classifying 1's and 2's using equal priors. What happens when the rules are adjusted

for unequal priors? "For groups of unequal sizes that tend to reflect relative population sizes (in an order sense), use of unequal priors will increase the hit rates for the larger groups and decrease the hit rates for the smaller groups" (Huberty, 1994, p. 112). When the increment in hits for the larger group exceeds the decrement in hits for the smaller group, the overall hit rate is higher. However, when the decrement in hits for the smaller group exceeds the increment in hits for the larger group, the overall hit rate is lower.

Consider data set #41 (Table 2), in which the correct (Pop+0) assumption about priors yields a lower hit rate than three of the incorrect assumptions (Equal, Pop+.15, Pop+.30). For this data set, Table 5 displays the average separate group and total hits for each of the six assumptions about priors. We can see how changes in separate group hits affect the results. Compared to the Equal assumption, for example, the Pop+0 assumption averages 19 <u>more</u> hits for Group 1 but 24 <u>fewer</u> hits for Group 2. Consequently, there are fewer total hits for the correct Pop+0 assumption than for the incorrect assumption of equal priors.

Insert Table 5 About Here

<u>Sample-Estimated Priors vs. Equal Priors</u>

Relative to the assumption of equal priors, the assumption of sample-estimated priors mirrored the Pop+0 pattern. When group sizes were similar, the Equal assumption was generally superior. When group sizes differed by 13% or more, the Sample assumption outperformed the Equal assumption.

Huberty (1994, p. 65) contends that sample estimated priors are inappropriate unless proportional sampling has been used. Results from the current study suggest that, perhaps even with proportional sampling, use of population priors may lead to a decrement in classification accuracy when group sizes are similar. Additional study is needed to confirm this interpretation.

References

Dolker, M., & Halperin, S. (1982). Comparing inverse, polar, and rectangle-wedge-tail FORTRAN routines for pseudo-random normal number generation. Educational and Psychological Measurement, 42, 223-226.

Huberty, C. J (1994). Applied discriminant analysis. New York: Wiley.

Knuth, D. E. (1969). The art of computer programming (Vol. 2: Seminumerical algorithms). Reading: MA: Addison-Wesley.

Lindeman, R. H., Merenda, P. F., & Gold, R. Z. (1980). Introduction to bivariate and multivariate analysis. Glencoe, IL: Scott, Foresman.

Marsaglia, G., MacLaren, D., & Bray, T. A. (1964). A fast procedure for generating random normal variables. Communications of the ACM, 7, 4-10.

Morris, J. D. (1975). A computer program to create a population with any desired centroid and covariance matrix. Educational and Psychological Measurement, 35, 707-710.

Morris, J. D., & Huberty, C. J (1987). Selecting a two-group classification weighting algorithm. Multivariate Behavioral Research, 22, 211-232.

Morrison, D. F. (1976). Multivariate statistical methods. New York: McGraw-Hill.

Park, S. K., & Miller, K. W. (1988). Random number generators: Good ones are hard to find. Communications of the ACM, 31, 1192-1201.

Rudolph, P. M., & Karson, M. (1988). The effect of unequal priors and unequal misclassification costs on MDA. Journal of Applied Statistics, 15, 69-81.

Tatsuoka, M. M. (1988). Multivariate analysis: Techniques for educational and psychological research (2nd ed.). New York: Macmillan.

Wichman, B. A., & Hill, I. D. (1982). An efficient and portable pseudo-random number generator. Algorithm AS 183. Applied Statistics, 311, 188-190.

Wilkinson, L. (1987). SYSTAT: The system for statistics. Evanston, IL: SYSTAT, Inc.

Table 1

Cross Validated Classification Performance (Portion of "Hits") for Six Assumptions About Prior Probabilities for Simulated Data Sets with Equal Group Sizes

#	Data Set Description	D	n_1	n_2	p	Assumption 1 Sample	2 Equal	3 Pop+0	4 Pop+.15	5 Pop+.30	6 Pop+.45
1	Fisher Data - Groups 1 & 3	12.63	1000	1000	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	Fisher Data - Groups 1 & 2	8.79	1000	1000	4	.9999	.9998	.9998	.9998	.9998	.9997
3	Bisbey Data - Groups 1 & 3	5.10	1000	1000	13	.9851	.9851	.9851	.9858	.9864	.9868
4	Fisher Data - Groups 2 & 3	4.24	1000	1000	4	.9792	.9794	.9794	.9804	.9809	.9809
5	Rulon Data - Groups 1 & 3	2.92	1000	1000	4	.9206	.9211	.9211	.9190	.9152	.9091
6	Bisbey Data - Groups 1 & 2	2.84	1000	1000	13	.9028	.9033	.9033	.9031	.9008	.8954
7	Bisbey Data - Groups 2 & 3	2.47	1000	1000	13	.8694	.8700	.8700	.8692	.8663	.8602
8	Talent Data - Groups 3 & 5	2.08	1000	1000	14	.8314	.8320	.8320	.8311	.8260	.8159
9	Demographic # 2 - Body Char	1.86	1000	1000	8	.8140	.8145	.8145	.8152	.8090	.7932
10	Rulon Data - Groups 2 & 3	1.86	1000	1000	4	.8182	.8188	.8188	.8158	.8062	.7891
11	Rulon Data - Groups 1 & 2	1.74	1000	1000	4	.8045	.8052	.8052	.8046	.7954	.7741
12	Talent Data - Groups 1 & 5	1.73	1000	1000	14	.7882	.7888	.7888	.7878	.7809	.7678
13	Demographic # 3 - Body Char	1.35	1000	1000	8	.7378	.7390	.7390	.7334	.7179	.6928
14	Talent Data - Groups 1 & 3	.92	1000	1000	14	.6357	.6365	.6365	.6286	.6135	.5924
15	Block Data - Groups 3 & 4	.85	1000	1000	4	.6566	.6621	.6621	.6733	.6590	.6275
16	Block Data - Groups 1 & 2	.85	1000	1000	4	.6469	.6522	.6522	.6285	.5927	.5563
17	Block Data - Groups 1 & 4	.81	1000	1000	4	.6348	.6392	.6392	.6145	.5811	.5487
18	Block Data - Groups 1 & 3	.75	1000	1000	4	.6255	.6314	.6314	.6037	.5677	.5365
19	Warncke Data - Groups 1 & 3	.67	1000	1000	10	.5935	.5955	.5955	.5945	.5840	.5657
20	Block Data - Groups 2 & 3	.66	1000	1000	4	.6061	.6095	.6095	.5870	.5589	.5335
21	Block Data - Groups 2 & 4	.50	1000	1000	4	.5657	.5692	.5692	.5602	.5433	.5269
22	Demographic # 1 - Body Char	.50	1000	1000	8	.5809	.5822	.5822	.5658	.5413	.5206
23	Warncke Data - Groups 1 & 2	.48	1000	1000	10	.5602	.5604	.5604	.5532	.5397	.5253
24	Warncke Data - Groups 2 & 3	.45	1000	1000	10	.5471	.5482	.5482	.5422	.5318	.5203

Note. The best performing assumption(s) are underlined ($p < .00001$).

12

Table 2

Cross Validated Classification Performance (Portion of "Hits") for Six Assumptions About Prior Probabilities for Simulated Data Sets with Group Sizes Proportional to Real Data Sets

#	Data Set Description	D	I	n_1	n_2	p	Assumption 1 Sample	2 Equal	3 Pop+0	4 Pop+.15	5 Pop+.30	6 Pop+.45
25	Fisher Data - Groups 1 & 3	12.63	100	1000	1000	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
26	Fisher Data - Groups 1 & 2	8.79	100	1000	1000	4	.9999	.9998	.9998	.9998	.9998	.9997
27	Bisbey Data - Groups 1 & 3	5.12	106	945	999	13	.9853	.9857	.9855	.9861	.9866	.9869
28	Fisher Data - Groups 2 & 3	4.24	100	1000	1000	4	.9792	.9794	.9794	.9804	.9809	.9809
29	Rulon Data - Groups 1 & 3	2.92	129	935	726	4	.9176	.9191	.9176	.9191	.9188	.9166
30	Bisbey Data - Groups 1 & 2	2.89	231	420	972	13	.9168	.9101	.9171	.9166	.9152	.9134
31	Bisbey Data - Groups 2 & 3	2.41	219	972	444	13	.8748	.8623	.8754	.8734	.8706	.8666
32	Talent Data - Groups 3 & 5	1.97	285	330	940	14	.8426	.8146	.8431	.8409	.8376	.8331
33	Demographic # 2 - Body Char	1.88	129	942	732	8	.8224	.8182	.8231	.8179	.8085	.7973
34	Rulon Data - Groups 2 & 3	1.87	141	930	660	4	.8251	.8202	.8263	.8227	.8151	.8043
35	Rulon Data - Groups 1 & 2	1.74	109	850	930	4	.8002	.8023	.8012	.8018	.7948	.7783
36	Talent Data - Groups 1 & 5	1.72	113	830	940	14	.7848	.7857	.7856	.7843	.7781	.7670
37	Demographic # 3 - Body Char	1.35	104	994	959	8	.7365	.7373	.7371	.7344	.7250	.7073
38	Talent Data - Groups 1 & 3	.89	252	996	396	14	.7109	.6433	.7114	.7018	.6903	.6765
39	Block Data - Groups 3 & 4	.85	100	1000	1000	4	.6566	.6621	.6621	.6733	.6590	.6275
40	Block Data - Groups 1 & 2	.85	105	1000	950	4	.6447	.6517	.6493	.6563	.6460	.6199
41	Block Data - Groups 1 & 4	.81	105	1000	950	4	.6327	.6398	.6371	.6462	.6386	.6132
42	Block Data - Groups 1 & 3	.74	103	1000	975	4	.6214	.6279	.6262	.6360	.6227	.5935
43	Warncke Data - Groups 1 & 3	.68	163	975	600	10	.6408	.6052	.6424	.6278	.6071	.5804
44	Block Data - Groups 2 & 3	.66	105	925	975	4	.6117	.6129	.6159	.5920	.5591	.5299
45	Block Data - Groups 2 & 4	.50	103	962	988	4	.5655	.5699	.5704	.5605	.5424	.5248
46	Demographic # 1 - Body Char	.50	104	959	994	8	.5836	.5833	.5845	.5674	.5409	.5182
47	Warncke Data - Groups 1 & 2	.48	138	975	705	10	.5801	.5661	.5811	.5702	.5520	.5285
48	Warncke Data - Groups 2 & 3	.44	118	987	840	10	.5527	.5488	.5521	.5454	.5342	.5188

Note. The best performing assumption(s) are underlined ($p < .00001$).

Table 3

Cross Validated Classification Performance (Portion of "Hits") for Six Assumptions About Prior Probabilities for Bootstrap Data Sets with Group Sizes Proportional to Real Data Sets

							Assumption					
							1	2	3	4	5	6
#	Data Set Description	D	I	n_1	n_2	p	Sample	Equal	Pop+0	Pop+.15	Pop+.30	Pop+.45
49	Fisher Data - Groups 1 & 3	13.97	100	50	50	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
50	Fisher Data - Groups 1 & 2	10.16	100	50	50	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
51	Bisbey Data - Groups 1 & 3	5.12	106	35	37	13	.9780	.9783	.9781	.9789	.9795	.9797
52	Fisher Data - Groups 2 & 3	3.77	100	50	50	4	.9298	.9306	.9306	.9293	.9245	.9149
53	Rulon Data - Groups 1 & 3	2.93	129	85	66	4	.9311	.9328	.9316	.9329	.9332	.9297
54	Bisbey Data - Groups 1 & 2	2.89	231	35	81	13	.9123	.9059	.9126	.9119	.9105	.9092
55	Bisbey Data - Groups 2 & 3	2.41	219	81	37	13	.8932	.8720	.8947	.8907	.8848	.8786
56	Talent Data - Groups 3 & 5	1.97	285	33	94	14	.8280	.8076	.8294	.8265	.8237	.8210
57	Demographic # 2 - Body Char	1.88	129	157	122	8	.8126	.8162	.8132	.8164	.8162	.8111
58	Rulon Data - Groups 2 & 3	1.87	141	93	66	4	.8258	.8225	.8267	.8255	.8170	.8025
59	Rulon Data - Groups 1 & 2	1.74	109	85	93	4	.8123	.8155	.8138	.8146	.7986	.7704
60	Talent Data - Groups 1 & 5	1.72	113	83	94	14	.7809	.7819	.7817	.7817	.7773	.7687
61	Demographic # 3 - Body Char	1.36	104	142	137	8	.7416	.7408	.7419	.7319	.7182	.7013
62	Talent Data - Groups 1 & 3	.89	252	83	33	14	.7106	.6620	.7110	.7050	.6980	.6889
63	Block Data - Groups 3 & 4	.85	100	38	38	4	.6658	.6749	.6749	.6632	.6267	.5900
64	Block Data - Groups 1 & 2	.84	105	40	38	4	.6647	.6787	.6764	.6734	.6456	.6197
65	Block Data - Groups 1 & 4	.81	105	40	38	4	.6125	.6184	.6163	.6254	.6216	.6092
66	Block Data - Groups 1 & 3	.74	103	40	39	4	.6299	.6405	.6390	.6382	.6173	.5875
67	Warncke Data - Groups 1 & 3	.69	163	65	40	10	.6485	.6159	.6511	.6390	.6182	.5872
68	Block Data - Groups 2 & 3	.64	105	37	39	4	.5926	.5992	.6011	.5752	.5478	.5305
69	Block Data - Groups 2 & 4	.52	103	37	38	4	.5998	.6073	.6079	.5945	.5635	.5331
70	Demographic # 1 - Body Char	.50	104	137	142	8	.5905	.5929	.5943	.5723	.5452	.5192
71	Warncke Data - Groups 1 & 2	.48	138	65	47	10	.5805	.5689	.5830	.5730	.5561	.5317
72	Warncke Data - Groups 2 & 3	.45	118	47	40	10	.5675	.5664	.5690	.5621	.5468	.5254

Note. The best performing assumption(s) are underlined ($p < .00001$).

Table 4

Cross Validated Classification Performance (Portion of "Hits") for Six Assumptions About Prior Probabilities for Simulated Data Sets with 900 1's and 100 2's

							Assumption					
							1	2	3	4	5	6
#	Data Set Description	D	I	n_1	n_2	p	Sample	Equal	Pop+0	Pop+.15	Pop+.30	Pop+.45
73	Fisher Data - Groups 1 & 2	12.96	900	900	100	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
74	Bisbey Data - Groups 1 & 3	4.74	900	900	100	13	.9873	.9826	.9876	.9875	.9874	.9873
75	Fisher Data - Groups 2 & 3	4.02	900	900	100	4	.9866	.9766	.9870	.9869	.9869	.9867
76	Rulon Data - Groups 1 & 3	3.06	900	900	100	4	.9628	.9280	.9633	.9628	.9618	.9607
77	Demographic # 2 - Body Char	2.06	900	900	100	8	.9285	.8462	.9287	.9277	.9266	.9250
78	Rulon Data - Groups 1 & 2	1.93	900	900	100	4	.9301	.8406	.9305	.9295	.9281	.9264
79	Talent Data - Groups 1 & 5	1.91	900	900	100	14	.9200	.8341	.9205	.9187	.9167	.9146
80	Demographic # 3 - Body Char	1.32	900	900	100	8	.9042	.7459	.9044	.9036	.9023	.9004
81	Block Data - Groups 3 & 4	1.23	900	900	100	4	.9169	.7770	.9167	.9162	.9152	.9141
82	Block Data - Groups 1 & 2	.75	900	900	100	4	.8926	.6498	.8928	.8905	.8875	.8844
83	Block Data - Groups 1 & 4	.73	900	900	100	4	.8933	.6517	.8938	.8916	.8888	.8853
84	Block Data - Groups 1 & 3	.68	900	900	100	4	.8948	.6242	.8953	.8937	.8917	.8892
85	Warncke Data - Groups 2 & 3	.43	900	900	100	10	.8868	.6764	.8867	.8829	.8787	.8742

Note. The best performing assumption(s) are underlined ($p < .00001$).

Table 5

Average Separate Group and Total Hits for Six Assumptions about Priors for Data Set #41

	Assumptions					
Average Hits	Sample	Equal	Pop+0	Pop+.15	Pop+.30	Pop+.45
Total	1233.820	1247.644	1242.416	1260.152	1245.284	1195.668
Group 1	633.240	616.400	635.124	526.100	415.496	303.612
Group 2	600.580	631.244	607.292	734.052	829.788	892.056